Predicting how many animals will be where: how to build, calibrate and evaluate individual-based models

Ecological Modelling 326 (2016) 113–123

Contents lists available at ScienceDirect

Ecological Modelling

journal homepage: www.elsevier.com/locate/ecolmodel

Predicting how many animals will be where: How to build, calibrate and evaluate individual-based models

Elske van der Vaart*, Alice S.A. Johnston, Richard M. Sibly

School of Biological Sciences, Harborne Building, University of Reading, Whiteknights, Reading RG6 6AS, Berkshire, United Kingdom

ARTICLE INFO

Article history:
Received 4 February 2015
Received in revised form 4 August 2015
Accepted 11 August 2015
Available online 4 September 2015

Keywords:
Energy budget
Individual-based models
Population dynamics
Approximate Bayesian Computation
Parameter estimation
Model selection

ABSTRACT

Individual-based models (IBMs) can simulate the actions of individual animals as they interact with one another and the landscape in which they live. When used in spatially-explicit landscapes IBMs can show how populations change over time in response to management actions. For instance, IBMs are being used to design strategies of conservation and of the exploitation of fisheries, and for assessing the effects on populations of major construction projects and of novel agricultural chemicals. In such real world contexts, it becomes especially important to build IBMs in a principled fashion, and to approach calibration and evaluation systematically. We argue that insights from physiological and behavioural ecology offer a recipe for building realistic models, and that Approximate Bayesian Computation (ABC) is a promising technique for the calibration and evaluation of IBMs.

IBMs are constructed primarily from knowledge about individuals. In ecological applications the relevant knowledge is found in physiological and behavioural ecology, and we approach these from an evolutionary perspective by taking into account how physiological and behavioural processes contribute to life histories, and how those life histories evolve. Evolutionary life history theory shows that, *other things being equal*, organisms should grow to sexual maturity as fast as possible, and then reproduce as fast as possible, while minimising *per capita* death rate. Physiological and behavioural ecology are largely built on these principles together with the laws of conservation of matter and energy. To complete construction of an IBM information is also needed on the effects of competitors, conspecifics and food scarcity; the maximum rates of ingestion, growth and reproduction, and life-history parameters.

Using this knowledge about physiological and behavioural processes provides a principled way to build IBMs, but model parameters vary between species and are often difficult to measure. A common solution is to manually compare model outputs with observations from real landscapes and so to obtain parameters which produce acceptable fits of model to data. However, this procedure can be convoluted and lead to over-calibrated and thus inflexible models. Many formal statistical techniques are unsuitable for use with IBMs, but we argue that ABC offers a potential way forward. It can be used to calibrate and compare complex stochastic models and to assess the uncertainty in their predictions. We describe methods used to implement ABC in an accessible way and illustrate them with examples and discussion of recent studies. Although much progress has been made, theoretical issues remain, and some of these are outlined and discussed.

1. Introduction

A major challenge in ecological modelling is to make reliable predictions about what will happen to real populations in real landscapes. In some ways this may seem a simple task—Newton solved similar problems in mechanics over 300 years ago. But animals and plants are not identical particles obeying simple mathematical laws, they make complex decisions based on their needs and perceived opportunities in their environments. Only with the advent of computing power has it become possible to simulate these processes with any degree of realism, and so to link the levels from individual organisms to populations of individuals. In this approach what happens to the population emerges from complex interactions between autonomous individuals and their environments, in the computer simulations as in life.

Models are always simplified representations of the real system, and so a trade-off is necessary between model complexity and realism (Evans et al., 2013). The different degrees of this

* Corresponding author. Tel.: +44118378547020.
E-mail addresses: e.e.vandervaart@reading.ac.uk (E. van der Vaart), a.s.a.johnston@pgr.reading.ac.uk (A.S.A. Johnston), r.m.sibly@reading.ac.uk (R.M. Sibly).

http://dx.doi.org/10.1016/j.ecolmodel.2015.08.012

trade-off are characterised by the different model types available. Differential equation models are typically used in simple assessments of unstructured population growth, whilst matrix models are essentially sets of linear difference equations which separate the population into classes (*e.g.* life-cycle stage) with class-specific life-history parameters (*e.g.* juvenile survival). Both approaches provide insight into general patterns of population growth in specified environmental conditions. They have the advantage that they can accept population-level data on birth and death rates, and they are often tractable using analytical methods. However they cannot easily accommodate autonomously acting individuals, and it is difficult to characterise the effects of location and habitat.

These high levels of detail can readily be incorporated into individual-based models (IBMs; also called agent-based models (ABMs)). In IBMs, the actions of unique individuals are simulated as they interact with one another and the landscape in which they live (DeAngelis and Mooij, 2005). Individuals can vary according to their state variables (*e.g.* age, sex, mass) whilst patches of mapped landscapes can be characterised by key ecological drivers (*e.g.* temperature, food, exposure to chemicals). The dynamics of populations in different environmental conditions then emerge from simulations of individuals' behaviours (Grimm and Railsback, 2005). Thus, where prediction is required about the fate of populations in different landscape scenarios, one way ahead is through IBMs (Stillman et al., 2015). Accordingly, IBMs are currently being used to design strategies of conservation and of the exploitation of fisheries, and for assessing the effects on populations of major construction projects and of novel agricultural chemicals (see, *e.g.*, Galic and Forbes, 2014; Hartman and Kitchell, 2008; Nabe-Nielsen et al., 2014; Stillman and Goss-Custard, 2010).

Although IBMs are powerful tools for ecological management, they also face major challenges. There may not be sufficient data available to build a realistic model, running IBMs may be computationally expensive, and run times may be prohibitively long. Furthermore attempts to represent multiple processes and interactions in IBMs can lead to models being over-parameterised, leading to reduced realism and an inability to extrapolate to other sites and/or time periods. Their predictions are then imposed rather than emergent (Grimm and Railsback, 2005; Martin et al., 2013). Because models are needed to forecast what happens in novel conditions, it is desirable that they be mechanistic in the sense that they accurately capture the underlying relationships between biological processes and environmental conditions.

In this paper we consider two particular problems: How to build ecological IBMs from first principles, and how to calibrate and evaluate them. When IBMs are built to predict the numbers and spatial distributions of animals, as is often the case in applied studies, we argue that insights from physiological and behavioural ecology offer a sound recipe for building realistic models. We also argue that model calibration and evaluation can be achieved using the new technique of Approximate Bayesian Computation (Beaumont, 2010). Thus the paper has two foci, which run in parallel but are not necessarily related to each other. Together they give our vision of "next generation ecological modelling", which is the focus of the special issue in which this paper appears. We try to produce concrete suggestions, but hope our readers will forgive us for not being able to fully describe the pros and cons of alternative approaches. This is partly for lack of space, but also in part because the new techniques we envisage are not yet fully developed or compared with alternatives, so informed comparisons and discussion are not yet possible. Our overarching aim is to be able to link the levels from individuals to populations in a transparent and credible fashion that is firmly rooted in biological knowledge.

2. Building IBMs from first principles

In this section we identify principles which may be used to build ecological IBMs and consider how to build such models using available biological knowledge. Our approach is partly based on Sibly et al. (2013) and is similar to the Dynamic Energy Budget approach (Kooijman, 2010; Martin et al., 2012). We then consider how population dynamics emerge from the simultaneous behaviours and interactions of individuals. At the end of the section we discuss some of the complications that arise in linking the levels from individuals to populations.

IBMs are based on knowledge about individuals, and the subdisciplines of biology that deal with individuals are physiological and behavioural ecology. These consider how physiological processes within individuals, and decisions made by individuals, contribute to life histories. Natural selection acts on life histories, favouring some at the expense of others, and this has ramifications for the evolution of physiologies and behaviour. So it is sensible to start by considering how life histories evolve.

The theory of life-history evolution is well established (see *e.g.*, Sibly, 2002; Stearns, 1992) and explains why organisms are expected to maximise Darwinian fitness and so to win out in the struggle for existence in the environment in which they evolved. In particular *other things being equal* organisms are expected to:

• Grow to sexual maturity as fast as possible	(Axiom 1)
• Reproduce as fast as possible	(Axiom 2)
• Minimise *per capita* death rate	(Axiom 3)

The phrase 'other things being equal' means that growth, reproduction and death rate are independent, *i.e.*, they do not trade off against each other. However this is not always the case, *e.g.*, growing faster may only be possible by taking risks, which may mean the death of the individual. In such cases organisms may trade off risk of death to increase their growth rate. Much attention has been given to the evolution of life histories that are subject to constraints imposed by life-history trade-offs (Sibly, 2002; Stearns, 1992). The predicted outcome of the evolutionary process in a constant environment is referred to as an *optimal strategy*, meaning the strategy that maximises Darwinian fitness subject to the imposed constraints. Constraints and opportunities differ among species, and this is one reason why species differ from each other. Incorporating trade-offs into IBMs can be straightforward; for instance, the increased mortality that comes with foraging in dangerous but rewarding places may be a direct result of encountering predators more often. Provided the different situations of different species are well-modelled, their different trade-offs should emerge automatically.

One major constraint to increasing Darwinian fitness stems from the availability of resources. The energy and nutrients needed to build animal bodies are derived from food, but food may be in limited supply. This imposes major constraints on behaviour and physiology as follows:

• Energy is conserved within individual bodies	(Axiom 4)

This means that the only energy available to power organisms is that which they derive from food or sunlight. Allocation of resources within bodies is similarly constrained:

• Matter is conserved within individual bodies	(Axiom 5)

This means that the only chemicals available to build organism bodies are those they derive from food.

Life-history theory is the foundation on which physiological and behavioural ecology are built. We now consider their relevant findings at the level of the individual.

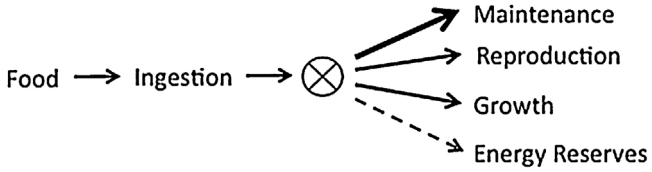

Fig. 1. Structure of the energy budget model. The thickness of solid arrows indicates priorities for allocation of energy obtained from food. Any energy remaining after these allocations enters the energy reserves.

2.1. Individuals

There are many complex and unresolved issues in ecological energetics that have to be reconciled with the need for simple representations that can be included in an IBM. The simplifications that follow represent our vision of how this can best be achieved—but bear in mind that ours is not the only possible approach and that others may prove superior as science progresses.

The main contribution of physiological ecology is understanding of the mechanisms of energy acquisition and expenditure by individuals, generally termed energy budgets. For modelling purposes the aim of energy budgets is to identify a generic specification for how individuals acquire and expend energy with sufficient realism but without unnecessary complexity. Our understanding of these processes, as derived from Glazier (2008), Karasov and Martinez del Rio (2007), Peters (1983) and Sibly and Calow (1986), is set out below, along with suggestions for how to model them mathematically.

We assume that an animal forages as necessary to supply its energy needs for maintenance, growth and reproduction. Maintenance here refers to the minimum energy requirements for survival, often taken as the basal metabolic rate (BMR). If there is sufficient energy intake, the animal allocates the energy obtained in the following order: maintenance, growth, reproduction, energy storage, until its energy stores reach an optimal level, as in Fig. 1 (Glazier, 2008; Karasov and Martinez del Rio, 2007; Peters, 1983; Sibly and Calow, 1986). This is a diagrammatic representation of Axioms 4 and 5, omitting faecal and excretory waste. The total available for allocation is limited by the amount the animal eats, so if more is allocated to one function, less is available for others. This follows from conservation of mass and energy (Axioms 4 and 5).

There is some but not much information as to how priorities change when there is not enough food (Glazier, 2008; Hou, 2014), but our view is that the priorities for maintenance and growth/reproduction remain the same until reserves fall to a critical threshold below which all is allocated to maintenance. Note however that one prominent theory, Dynamic Energy Budget (DEB) theory, makes a different assumption, that throughout life a constant fraction of input is allocated to maintenance and growth, with the rest going in juveniles to maturation and in adults to reproduction, (the "kappa rule", Kooijman, 2010). Calculations are generally in units of energy per unit time, *e.g.* watts, even though acquisition and allocation of many specific nutrients subscribe to the same principles (see *e.g.*, Kaspari, 2012).

Food acquisition and digestion. According to the principles of optimal foraging, food resources are generally chosen from those available according to the net rate at which they provide energy per unit time (Davies et al., 2012, see also Section 2.3). Thus:

- When foods vary in energy yield per unit time after allowing for energy costs of foraging, the animal selects the most profitable (Axiom 6)

Food resources generally vary both temporally and spatially. Variation in food density affects the rate of ingestion of food up to an asymptote, the form of this relationship being known as a 'functional response', and generally this is modelled as a two-parameter Holling type 2 response (Holling, 1959), which often approximates what is observed in nature (Begon et al., 2006; Krebs, 2009; Ricklefs and Miller, 2000). The Holling type 2 functional response may be written:

$$\text{Ingestion rate} = IG_m \times \frac{(\text{food density})}{(\text{food density} + h)} \quad (1)$$

where IG_m is the maximum ingestion rate in g or J per unit time, and h is a constant which shows how quickly the response curve reaches its maximum as density increases. Maximum ingestion rates generally scale allometrically with body mass and temperature (Clauss et al., 2007; Peters, 1983).

After ingestion food is processed by the digestive system and a proportion becomes available for allocation to the various functions shown in Fig. 1. This proportion is called assimilation efficiency, defined as: (energy obtained by digestion)/(energy ingested as food). Assimilation efficiency depends on diet and averages around 50–60% (Peters, 1983) and appears not to vary with body mass (Hendriks, 1999). However, assimilation efficiency varies widely between diets. Whereas flesh and seeds may be upwards of 80% assimilated, this falls to 40–70% for young vegetation, and lower for mature vegetation and wood (Peters, 1983). Hendriks (1999) gives the assimilation efficiencies of detritivores, herbivores and granivores/carnivores as around 20%, 40% and 80%, respectively. Assimilated energy is available for distribution to maintenance, growth, reproduction and energy reserves as described in the following sections.

Maintenance and survival: Energy for maintenance is roughly equivalent to BMR, and the dependence of BMR on body mass, M, and body temperature T, measured in Kelvins (=°C + 273.15), can be approximated as:

$$\text{Metabolic rate} = B_0 M^{\alpha} e^{-E/kT} \quad (2)$$

where B_0 is a constant of proportionality, α is a scaling coefficient, ~3/4 (Glazier, 2005; Moses et al., 2008; Peters, 1983), E is activation energy in eV, ~0.65, κ is Boltzmann's constant, and the exponential term is sometimes referred to as the Arrhenius function (Brown and Sibly, 2012; Peters, 1983). Energy allocated to maintenance fuels the basic processes of life essential for survival and these have first call on energy obtained from feeding, and on an animal's energy reserves when food is unavailable. Energy is allocated to maintenance as long as energy is left in the reserves. For modelling purposes the animal may be considered dead when the reserves are exhausted.

Growth: If energy is available after the costs of maintenance have been paid, juveniles allocate energy to somatic growth. The energy costs of growth, per gram of flesh synthesised, are fairly well known (wet flesh contains 7 around kJ/g, and the energy cost of synthesis is around 7 kJ/g for homeotherms and 3.6 kJ/g for poikilotherms (Sibly et al., 2013)). However there are limits to the rates at which animals can grow and these change as the animal grows. How these limits change with body mass has been variously modelled, but the resulting growth curves are very similar (Kerkhoff, 2012). A widely used model of growth rate under optimal conditions in relation to bodymass M at time t is

$$\frac{dm}{dt} = r_B \left({M_m}^{1/3} M^{2/3} - M \right) \quad (3)$$

where M_m denotes maximum body mass and the parameter r_B is the coefficient in the von Bertalanffy equation, which can be obtained from data recording increase of body length or mass with age in ideal conditions. There has been controversy as to the mechanistic underpinnings of Eqs. (2) and (3) and the exact values of their exponents (see, *e.g.*, Kerkhoff, 2012; Price et al., 2012); we do not endorse one model over others but suggest these equations as commonly used ways of describing the relationships. Eq. (3) shows how the maximum rate at which resources can be allocated

to growth changes as the juvenile increases in mass. Another factor affecting growth rate in ectotherms is body temperature. The effect of temperature on growth is given by the Arrhenius function referred to in Eq. (2). The case of continued growth after first reproduction is more complicated and is discussed in Sibly et al. (2013). If any energy remains after paying the costs of maintenance and growth and perhaps reproduction, it goes into energy reserves.

Reproduction: Reproduction does not occur until the animal has attained a certain size and assembled the bodily structures necessary for reproduction. These structures (*e.g.*, gonads, oviduct, uterus) themselves require resources and some models account for this explicitly (*e.g.*, Kooijman, 2010) but this may not be necessary provided a minimum size (or age) of reproduction is included.

Reproduction, like growth, requires that molecules be precisely assembled in appropriate order, and this imposes limits on the rate at which new flesh can be synthesised in developing embryos. The maximum rates of production are implicit in the allometric coefficients for numbers and sizes of offspring, and these are available with age at maturity for many species in the literature (Sibly et al., 2013). The energy cost of synthesising flesh for reproduction is the same as for growth.

Food supply and in some species temperature affect when an animal reaches the size required for reproduction. For determinate growers that size would be adult size. However, while this approach may suffice for many vertebrates, some invertebrates respond to food shortage/stress in more complex ways, by decreasing size of first reproduction and clutch size, and in some species by increasing neonate mass. Some of these invertebrates are indeterminate growers, and these are discussed in Sibly et al. (2013).

Energy reserves: Energy reserves in terrestrial vertebrates are stored mainly as fat in adipose tissue, containing 39 kJ/g, or as carbohydrates in the liver (18 kJ/g) (Sibly et al., 2013). These reserves allow the animal to maintain its functions during temporary periods of starvation. If energy input from food exceeds the requirements of maintenance, growth and reproduction, then any excess is stored in the animal's energy reserves. Conversely reserves are used to supply energy requirements if the supply from feeding is inadequate. There are costs to energy storage and the total cost of synthesising and storing one gram of fat is about 54 kJ. Despite the attractions of fat some animals use other fuels, *e.g.*, sessile marine animals, for which carrying extra weight is not costly, use glycogen, while earthworms and flatworms use protein and degrow when starving.

Surplus energy from food is not added to reserves indefinitely. Instead animals stop eating once reserves reach a certain level, presumably corresponding to an optimum compromise between the benefits of being able to survive a hunger gap and the costs of carrying extra weight, *e.g.*, reduced ability to escape from predators (Gosler et al., 1995; Lind et al., 2010; Witter and Cuthill, 1993). The optimum will vary with time and place, and prior to migration animals may accumulate a fat store of 25–50% of body mass (Peters, 1983; Pond, 1978). While optimum values of energy reserves cannot be predicted *a priori*, information on natural fat content exists for many species (see *e.g.*, Pond, 1978). Relative to energy expenditure larger mammals carry more body fat than smaller ones (fat = $75 \times M^{1.19}$, fat in g and M in kg, Lindstedt and Schaeffer, 2002), and so can survive substantially longer periods of starvation.

In this section we have considered how individuals obtain and process food, and how they use it to fuel maintenance, growth and reproduction. In our exposition survivorship is maximised and growth and reproduction occur as fast as possible in the absence of trade-offs, in accordance with Axioms 1–3. However where trade-offs exist the optimal strategy may not be predictable, as with the optimal level of energy reserves discussed above. The assembly of individuals in the modelled landscape constitutes a population, and we turn next to how such populations can be studied.

2.2. Populations

Here we consider the ways in which populations are affected by features of their modelled environments, and how environmental effects can be identified. In the first place we note that individuals do not act completely independently of each other. Instead, the actions of individuals both influence and are influenced by the actions of other individuals. For instance, classical ecological processes such as habitat selection, competition, predator–prey interactions and dispersal all depend on the physiological and behavioural interactions between individuals and other individuals and the landscape (*e.g.* Sih et al., 2012). It is one of the strengths of IBMs that population dynamics emerge from explicit simulations of these processes. However to understand the causes of the patterns of emergent population dynamics generally requires further work.

As an example consider Dalkvist et al.'s (2011) study of vole population dynamics in Fennoscandia. Vole population dynamics vary systematically from regular cycles in the north to stable populations in the south. The reasons are believed to include properties of the voles' predators and habitat fragmentation, but these also vary from north to south, and their effects are hard to distinguish in field experiments. However both can be manipulated in IBMs. Dalkvist et al. (2011) showed by experimentally manipulation of IBM landscapes that both habitat fragmentation and the presence of specialist predators are necessary for the occurrence of population cycles, and the properties of the predators and the habitats, together with those of the voles, jointly determine vole cycle length and amplitude.

The effects of habitat fragmentation on the long-term persistence of wild animal populations are also important to wildlife managers and conservation biologists, but as with the voles it is rarely feasible to undertake field experiments to establish the effects of habitat fragmentation. In an attempt to obtain some general insight Nabe-Nielsen et al. (2010) used IBMs to look at the effects on skylarks, voles, and particular ground beetles and spiders, of progressively fragmenting a real 10×10 km Danish landscape. The most important result was that the arrangement of habitat patches and the presence of corridors had a large effect on the population dynamics of species whose local success depends on the surrounding terrain. Similarly Liu et al. (2013) showed how the adverse effects of pesticides on wood mouse populations could be reduced by the addition of favourable hedgerow habitats. While these results may be intuitive, the use of IBMs allows predictions to be made as to what will happen if specified modifications are made to the landscape, for instance by introducing corridors such as unmanaged grassland for voles, or vegetated field boundaries for beetles.

These examples illustrate how IBMs can be used to predict management effects on populations living in real landscapes. However to achieve realistic predictions it is not enough just to build an IBM and show the emergent population dynamics. Further understanding is generally needed to establish what causes particular population phenomena, such as cycles. If the underlying causes are accurately identified, the population predictions should be realistic. However, establishing realism is always difficult. We turn next to some of the problems that may arise when using IBMs to make links from individuals to populations.

2.3. Complications in linking the levels from individuals to populations

Understanding of the causes of population dynamics requires accurate models of how individuals behave and allocate resources. However, lack of knowledge and data at both the levels of

individuals and populations may place limits on what can be achieved. Some of the major complications are discussed in the following section.

First, when modelling decision making at the individual level, *e.g.* about what to eat, it is often assumed that animals optimise fitness. Optimal decision models have been very successful in understanding animal behaviour, however problems remain (Davies et al., 2012). Since effects on Darwinian fitness can generally not be measured directly, a surrogate such as rate of obtaining energy is used instead. However, the surrogate may not accurately reflect effects on Darwinian fitness. For instance we have until now been assuming that the essential requirement of animals is for energy, but other nutrients may sometimes be limiting. For example, some fledgling birds require insects rather than grain to grow and develop properly, and the diet of herbivores may lack essential salts, forcing animals to seek salt licks. In such cases the independent needs for nutrients and energy would need to be modelled separately, though in principle this can still be achieved within a framework of maximising Darwinian fitness (Simpson et al., 2004).

Second, our assumptions about what would be optimal decision making may be wrong if we do not correctly identify an animal's physiological limitations, such as the time required to crack a prey's defences before it can be consumed.

Third, it may be necessary to incorporate the fact that individuals do not have perfect information about their environment, and instead need to rely on sampling and memory. Such insights can be implemented into IBMs fairly easily, in contrast to other types of modelling approaches. In an IBM of woodpigeon flocks, Kulakowska et al. (2014) showed that a model in which individuals forage optimally did not adequately fit data from radio-tracking studies and other data from a 40-year study of the distribution of birds between crops. To obtain adequate fits of the available data it was necessary to allow that individuals had imperfect knowledge of their environment, and had to rely instead on memories of previous experiences.

Finally, complex social interactions may be a major source of complications in modelling. For instance, many animals live in groups and within these groups the distribution of resources between individuals may be affected by a dominance hierarchy and/or by nepotism (helping relatives). Moreover, there may be conflicts for resources between groups, indeed this is seemingly inevitable given that most populations are food limited (Sinclair, 1989). In inter-group competitions some groups may prosper and grow and perhaps eventually split into subgroups when some individuals would fare better with fewer companions. Many IBMs already simulate social interactions, including dominance hierarchies (*e.g.* Puga-Gonzalez et al., 2009; Evers et al., 2012) and group movement (*e.g.* Petit et al., 2009), but these tend to be theoretical models looking at fundamental questions. Incorporating social interactions into practical, prediction-based IBMs of real populations in real landscapes remains an open challenge. There are myriad potential complications and variations stemming from social interactions, reflecting the diversity of the natural world.

We conclude from this section that IBMs should incorporate insights from physiological and behavioural ecology, since these represent the current state of scientific knowledge. However complications such as those outlined above show how important it is to realise that this approach alone does not ensure realism. Realism is evaluated by assessing how well the models outputs match independent data at both the individual and population level (Grimm and Railsback, 2012). Methods for achieving this are described and discussed next.

3. Calibrating and evaluating models

One challenge that arises when attempting to build realistic IBMs is the need to estimate the values of model parameters. For instance, even the simplest energy budget model contains a fair number of parameters. Although some of these, like the energy cost of synthesis, E_s, are fairly well-known, others, like the maximum ingestion rate IG_m, are highly species specific and can be difficult to measure. A related problem occurs when trying to design an IBM's *structure*. A model should be as simple as possible, but no simpler, and it is not always clear what line divides the two. Is it necessary to simulate prey types individually? Or seasonal changes in the weather? Or the dynamics of social interaction? Even taking the insights from physiological and behavioural ecology into account, deciding which mechanisms to include in any given IBM can be surprisingly difficult.

Current best practice for ecological IBMs, both for parameter estimation and for model choice, is known as 'pattern-oriented modelling' or POM (Grimm and Railsback, 2005, 2012). The basic idea behind POM is to try to simultaneously fit multiple, ecologically-relevant patterns, preferably at different levels of biological organisation. Essentially, POM is a protocol, where each 'pattern' serves as a filter that can either suggest or reject particular model configurations. As an example, Topping et al. (2012) used POM to calibrate their existing field vole model. They defined sets of patterns relating to population structure, habitat use, dispersal distance and predator/prey cycling, and this prompted adjustments of their original model. Specifically they found they needed to explicitly simulate live-traps to obtain outputs comparable to the empirical data, and that additional parameters were necessary to capture variation in vole density across habitats.

Although POM works well, it is 'experimental and largely based on experience', as Topping et al. (2012) acknowledge. Moreover, for POM to be useful to decision makers, a more quantitative approach is needed to evaluate the relative strengths of different models in making predictions for specific purposes. We believe that Approximate Bayesian Computation, or ABC, can complement POM in the IBM modelling cycle: It preserves the basic ideas of the method, while at the same time making it more transparent and statistically rigorous.

ABC is a method for quantifying the support that a given set of data lends to particular model choices. This is achieved by computing the probabilities of both *parameter values* and *model alternatives* given the *data*. What makes ABC 'Bayesian' is that it is about updating degrees of belief. One starts with *prior probabilities* for all parameters and model versions and ends up with *posterior probabilities*. What makes ABC 'approximate' is that it does not require deriving these probabilities analytically, which in not generally possible for IBMs; instead, they are approximated through *simulation*. This makes ABC one of very few methods of model analysis that will actually work with IBMs; however, the use of ABC for IBMs is still in its infancy.

The birth of ABC is often traced to a series of papers published at the turn of the century (Beaumont et al., 2002; Pritchard et al., 1999; Tavaré et al., 1997), all motivated by problems in population genetics. Since then, the majority of the ABC literature has been written for this audience, or for statisticians. This creates a significant entry barrier for individual-based modelers (though see Bertorelle et al., 2010; Csilléry et al., 2010; Hartig et al., 2011, for accessible reviews), who are often unfamiliar with the relevant language and examples. Despite this, several authors have noted the potential that ABC offers for IBMs (Beaumont, 2010; Thiele et al., 2014; Topping et al., 2012) and there are now a few successful applications (Hartig et al., 2014; Martínez et al., 2011; van der Vaart et al., 2015a). ABC has been developed in sophisticated variants (see Section 3.4) but here we only describe the simplest approach, sometimes termed "rejection-ABC" (Beaumont, 2010). Our aim is to provide a gentle introduction to ABC with an example of ABC in practice, and some new results and discussion.

3.1. Estimating parameters with rejection-ABC

So, how does rejection-ABC actually work? The basic 'recipe' for doing parameter estimation with rejection-ABC is given by the following procedure:

(1) Select the empirical data that the IBM should fit and set up the IBM accordingly.
(2) Define *prior distributions* for all of the model's parameters.
(3) Run the IBM *e.g.* 10^5 times, using random samples from its prior distributions.
(4) Accept the *e.g.* 100 runs which provide model outputs which best fit the empirical data.
(5) Analyse the accepted parameters to obtain approximate *posterior distributions*.
(6) Check the IBM's fit using the accepted parameters—the *posterior predictive check*.
(7) Check the accuracy of the estimation processes using *e.g. cross-validation* and *coverage plots*.

Step 1 is to select the relevant empirical data: What patterns must the IBM replicate to be considered 'fit for purpose'? As in POM, these patterns can be at different levels of organisation – *e.g.*, some may be at the individual level, some at the population level – but unlike POM, they must all be expressed numerically. If the available empirical data is too detailed to be effectively compared to the model output, it may be necessary to summarise it. For example, instead of using the location data of every individual at every timestep, it may be better to use the average path length, or the percentage of time spent in every habitat, as *summary statistics*. The IBM must then be set up so that it replicates the conditions that produced the empirical data and produces matching outputs.

Step 2 is to define *prior distributions* for all of the model's parameters. For example, within what range of values are they likely to lie? These prior distributions can take on any shape that accurately reflects what is actually known—for instance, wide, uniform priors when very little information is available, and tight, normally distributed ones around existing values if these are likely to be correct.

In Step 3, the IBM is run anywhere from thousands to millions of times, using independent, random samples from its priors, resulting in *e.g.* 10^5 sets of simulated 'summary statistics'–model outputs summarised the same way as the empirical data was in Step 1. To achieve Step 4 we need a measure of distance between data points and the corresponding model outputs. How is this distance to be defined? A straightforward method is to calculate the Euclidean distance ρ between the model output of run i and the empirical data points $D_{j=1,2\ldots n}$ using the equation:

$$\rho(m_i, D) = \sqrt{\sum_j \left(\frac{m_{i,j} - D_j}{sd\,(m_j)} \right)^2} \quad (4)$$

where $m_{i,j}$ is run i's output for summary statistic j, D_j is the empirical data for summary statistic j, and $sd\,(m_j)$ is the standard deviation of summary statistic j in all model runs (Beaumont, 2010). Here $sd\,(m_j)$ is a scaling factor used to normalise the scales of the various summary statistics; for instance, body masses may be in tens of grams while eggs laid per week are in single figures. If the differences between the model outputs and the empirical data were not appropriately scaled, the distance calculations would be dominated by the body masses, because of the choice of units used to measure them. In sum, ρ measures the discrepancy between the model outputs and the corresponding empirical data points

In Step 4, some of the runs that minimise ρ are accepted as 'close enough' to the empirical data. The number of runs to accept may be determined pragmatically. At least 100 or so are needed to generate reliable posterior frequency distributions of the parameter values, though if the model is very stochastic more might be needed. On the other hand it is attractive to use as few as possible so that those used give good fits of model outputs to the data. In our experience accepting 100 achieves a pragmatic compromise between two conflicting desiderata.

In Step 5, the distribution of parameter values in the accepted runs is analysed, and this yields an approximate *posterior distribution* for each parameter. In addition, a *point estimate* is often computed, some summary of the posterior distribution which reflects ABC's 'best guess'—this may be the value that gave the best-fitting run, or the median of the accepted values. Step 6 is to do a *posterior predictive check*—to sample the accepted runs randomly, and to use their parameter values to re-run the IBM, in order to investigate how well they cause the IBM to fit the data.

Finally, Step 7 may be used as a form of quality control. Two useful diagnostic procedures are *cross-validation* (Csillery et al., 2012) and *coverage* (Prangle et al., 2013). The ideas behind both are similar. Both use model runs that have already been performed in Step 3. Some of these runs are set aside as "pseudo-data" and then the remaining runs are used to check whether ABC can correctly estimate the parameter values that generated them. One difference between cross-validation and coverage is that the former looks at ABC's ability to produce correct point estimates of parameter values, while the latter looks at the accuracy of the posterior distributions. The results of both cross-validation and coverage are plotted in diagnostic plots, allowing the modeller to diagnose potential problems in ABC's estimation procedures.

3.2. Model comparison with rejection-ABC

If instead of estimating a model's parameters the goal is to compare structurally different models, it is necessary to do *model selection*. The procedure for this is similar to parameter estimation, except that *each* model must be run *e.g.* 10^5 times using random samples from its priors. In Step 5, the ratio of accepted models gives their probability given the data. If, for instance, 80 copies of model A were accepted and 20 copies of model B, the empirical data favours model A over model B by a factor of 80/20 = 4. This factor is known as the *Bayes factor* $B_{A,B}$ and expresses the degree to which the empirical data favours model A over model B. Some suggest that a Bayes factor of 1–3 counts as 'barely worth mentioning', 3–10 counts as 'substantial evidence', 10 to 100 as 'strong evidence' and >100 as 'decisive' (Kass and Raftery, 1995).

What makes rejection-ABC model selection especially attractive is that it automatically corrects for differences in model complexity, provided each model is run equally often (Beaumont, 2010). This is because the more parameters a model has, the more sparsely the 'correct' parameter settings will be sampled; the consequence is that the more complex model will be accepted more often only if the additional parameters contribute enough additional explanatory power. This is a very useful feature when it comes to comparing IBMs, whose degrees of freedom can be difficult to determine.

3.3. A worked example

To get a better sense of what rejection-ABC can do for IBMs, it is useful to discuss an example. Previously, we have used the rejection-ABC procedure outlined above to calibrate a 14-parameter energy budget IBM of the earthworm *Eisenia fetida* (van der Vaart et al., 2015a). The energy budget broadly follows that outlined in the previous section of this paper, where fundamental principles of physiological ecology are modelled as a set of metabolic equations. Then, interactions between individuals and their landscape in the IBM lead to emergent population patterns.

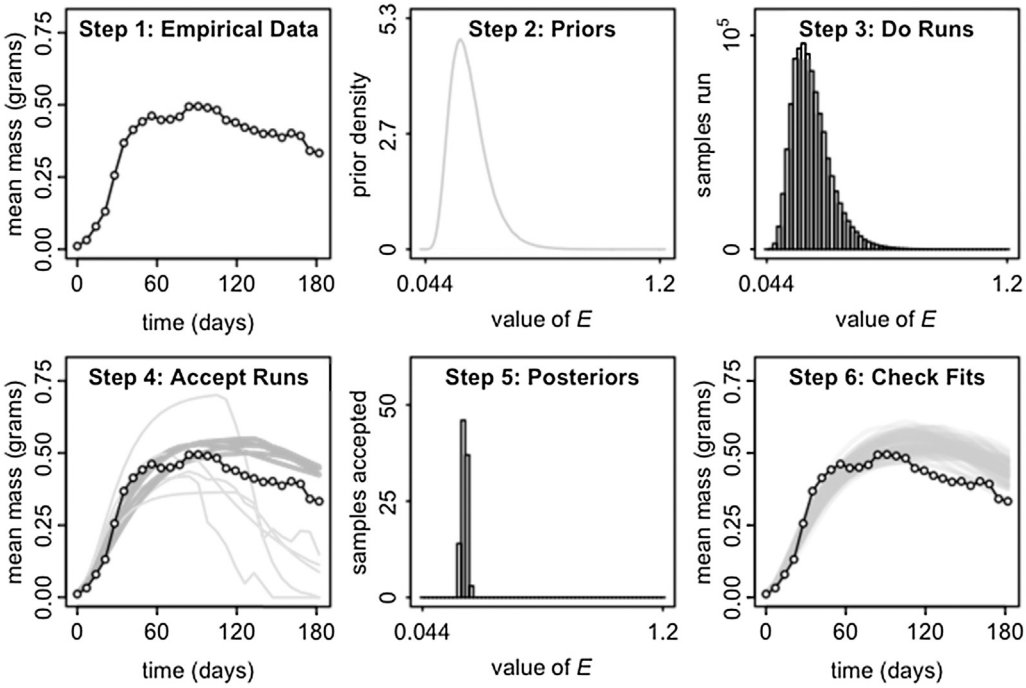

Fig. 2. Summary of an example ABC analysis. Step 1 is to select the relevant empirical data. We used mean body masses and cocoon totals from Gunadi et al. (2002), Gunadi and Edwards (2003) and Reinecke and Viljoen (1990); mean body masses from Gunadi et al. (2002) are shown. Step 2 is to choose prior distributions for all 14 parameters; given here is the lognormal prior for *E*, the activation energy, centred around a value previously taken from the literature. In Step 3 the model is run many times, using random samples from the priors; we used one million runs in total. Step 4 is to accept the runs closest to the empirical data; thick dark grey lines show 5 out of 100 accepted runs, thin light grey lines show 5 example rejected runs. Step 5 is to analyse the posteriors of all parameters; shown again is *E*, which was significantly narrowed. Finally, Step 6 is to do a posterior check, verifying how well the model fits when re-run with the accepted parameters.

In the following sections we provide two examples using rejection-ABC to calibrate and evaluate models. First, we provide a summary of previous work (van der Vaart et al., 2015a), where we used rejection-ABC to calibrate the model according to individual-level patterns of growth and reproduction. Second, we repeat the process using population-level data to explore how the individual and population levels can be linked. All simulation results, the earthworm IBMs and the ABC code have been deposited in a figshare repository (van der Vaart et al., 2015b,c), along with a brief guide to their use.

3.3.1. Individuals

In previous work we fitted an energy-budget model to empirical data consisting of measurements of individual growth curves and cocoon production of laboratory-kept earthworms (Step 1). The prior distributions were lognormals, with medians equal to values previously calculated from the literature (Step 2). One million simulation runs were made (Step 3), all with unique parameter combinations, of which 100 were accepted (Step 4). Interestingly, we found that only seven of the model's fourteen parameters were significantly narrowed (Step 5) but that the IBM nevertheless fitted the empirical data rather well (Step 6). The whole process is briefly summarised in Fig. 2.

When we investigated *why* only seven of the model's parameters were narrowed, we found that in part this was due to correlations between parameters: In the accepted runs, for instance, the value of r_m, the maximum energy allocation to reproduction, was positively correlated with M_c, the mass of cocoons. This means that the empirical data was not sufficiently detailed to tell the difference between earthworms spending a lot of energy making heavy cocoons, and little energy making light cocoons. On the one hand, this suggests that further development of the earthworm IBM would be aided by adding in a data set that includes cocoon masses. On the other hand, it suggests that perhaps the empirical data that is already available could be fit by a simpler model.

To illustrate the power of rejection-ABC to compare models and to see whether a simpler model would fit the data equally well, we then built a simpler model, as follows. We removed the earthworms' movements, the effect of food density on food intake, and most of the model's energy budget dynamics. In the resulting simpler model individuals kept growing and reproducing maximally every day that they had any food to eat. When there was no food they shrank sufficiently to cover their maintenance costs. Using rejection-ABC model selection, we contrasted this simpler model with the full model. The result was that the simpler model was much less successful in fitting the data. We ran each model one million times and accepted the 200 that fitted the data best. Of these 200, two were produced by the simple model and 198 by the full model, leading to a Bayes factor $B_{\text{full, simple}}$ of 99 because the full model was accepted 99 times as often as the simple model. This is strong evidence that the full model fits the data better than the simpler model, giving confidence in the inclusion of the energy budget in the full earthworm IBM.

3.3.2. Populations

In the above approach we used rejection-ABC to investigate the parameterisation and structure of the earthworm IBM applied to data on individuals maintained in the laboratory. However data are also available, albeit of lower quality, from a population field study. These would be attractive to include in the analysis since it would allow us to truly link the levels from individuals to populations, one of the themes of this paper. In a field study Monroy et al. (2006) counted and weighed the earthworms in a population in a Spanish manure heap every season for a year, and Johnston et al. (2014) estimated the likely corresponding rainfalls and temperatures. As a first step towards including this new population data set, we took the posterior parameter distributions from our earlier

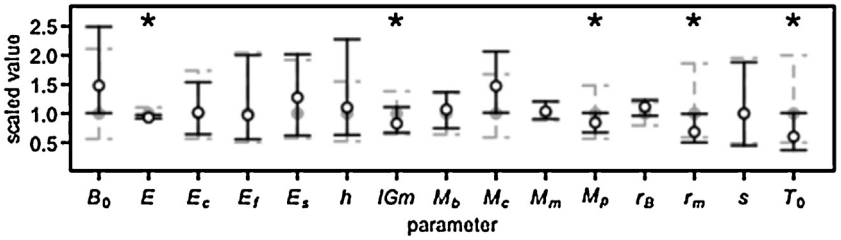

Fig. 3. Distributions of parameter values. Grey lines show the distributions of the priors, which were the result of fitting the earthworm IBM to laboratory data as described in Fig. 2; black lines show the distributions of the posteriors after fitting to the population study of Monroy et al. (2006). Circles represent medians, whiskers 95% credible intervals. Asterisks mark significant narrowing. All parameter values were scaled by dividing by the median of the corresponding prior. From left to right, the parameters listed are the taxon-specific normalisation constant, activation energy, energy cost of tissue, energy from food, energy cost of synthesis, half saturation coefficient, maximum ingestion rate, mass at birth, mass of cocoon, maximum asymptotic mass, mass at sexual maturity, growth constant, maximum energy to reproduction, movement speed and the cocoon hatching time, respectively. See van der Vaart et al. (2015a) for more information.

analysis and used them as the prior distributions for a new analysis. The results of rejection-ABC suggested some further narrowing of a subset of parameter values (Fig. 3), and obtained good fits in posterior checks (Fig. 4), with the exception of overestimating the number of juveniles in summer.

One might expect parameter values to differ noticeably when different datasets are used as a basis for their estimation. However, rejection-ABC parameter estimates when population level data is used are little changed compared to those estimated from individual level data (Fig. 3). This suggests that the model structure is sufficiently mechanistic to link the levels from individuals to populations. The mechanistic link is derived from physiological ecology as outlined in the first half of this paper. Our hope is that IBMs in which individuals have their own energy budgets will produce reliable predictions where individual life histories vary as a result of environmental variation in *e.g.* food availability or temperature.

3.4. Discussion

While rejection-ABC as outlined above is conceptually straightforward, its use in practice requires the availability of suitable hardware and software. It takes a desktop PC about half a second to run the earthworm IBM through the four laboratory experiments of Fig. 2, but the population field study takes a minute. To do a million runs sequentially would take over two years. Many IBMs, especially those incorporating large-scale, realistic geographic data will take even longer. In these situations, having access to a large cluster or supercomputer is essential. Fortunately, because the simulations required by rejection-ABC are completely independent, it is very easy to run them in parallel; it is as simple as starting as many copies of the IBM as there are computer cores available. We run our simulations on ARCHER, the UK's national supercomputer, using up to 50,000 cores at a time, but even on a desktop PC with four or six cores, the speed gained by parallelisation is considerable.

The easiest way to run any given IBM in parallel will depend on the programming language in which it is written. For NetLogo, a software platform designed specifically for developing IBMs (Wilensky, 1999), there are at least two useful tools available. First, NetLogo itself comes with BehaviorSpace, a built-in functionality that allows the user to run simulations in parallel from a drop-down menu. It can only be used for investigating uniform priors, spaced on a grid, but is very easy to use. A more powerful option

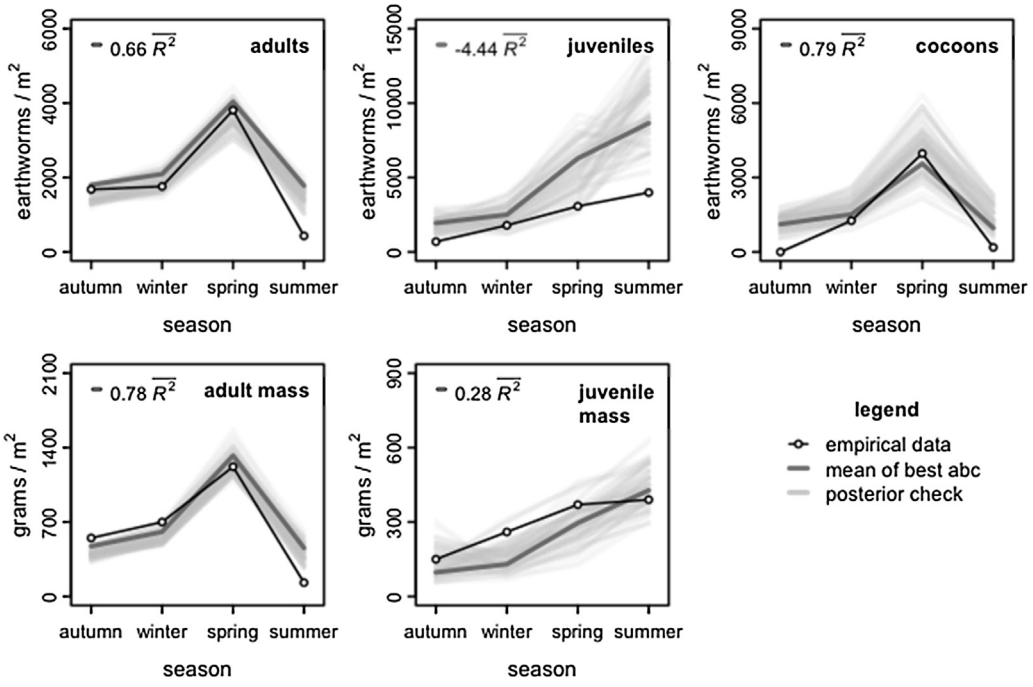

Fig. 4. Fits of the earthworm IBM to the empirical data. Results after calibrating with the empirical data from Monroy et al.'s (2006) field study. The open circles are the empirical data, and the semi-transparent grey lines are the 'posterior predictive checks', *i.e.*, the output of 100 new simulations using random samples from the accepted runs. Thick grey lines represent the mean of 100 simulations using ABC's best-fitting parameter set, and R^2 is the mean proportion of variance explained by these best-fitting simulations, a measure of goodness of fit.

is to use *R*, statistical software that, like NetLogo, is freely available for all operating systems. *R* comes with many built-in distributions to draw priors from, and the packages *RNetLogo* (Thiele et al., 2012, 2014) and *parallel* together provide a means of performing NetLogo runs in parallel.

The rejection-ABC analysis – Steps 3 and 4 of our rejection-ABC recipe above – can also be handled well by *R*. The *R* package *abc*, for example, takes empirical data, a spreadsheet of priors and a spreadsheet of results as its inputs, and produces as outputs the posterior distributions of Step 5 as well as the cross-validation diagnostics of Step 7. Other relevant *R* packages are listed by Thiele et al. (2014). A future objective of ours is to release an *R* package which will automate the rejection-ABC process for NetLogo models from start to finish; a beta version is available upon request.

There are two common ways in which the basic rejection-ABC algorithm introduced in this paper is sometimes modified. First, it may be possible to sample a model's priors more efficiently in Step 3, using either MCMC-ABC (Marjoram et al., 2003) or SMC-ABC (Sisson et al., 2007; Toni et al., 2009). MCMC-ABC bases each subsequent run of the model on the previous one, and gradually moves towards an estimate of the full posterior distribution. SMC-ABC starts a set of simulations in parallel, sampling randomly from a model's priors, but then gradually lowers the acceptance rate, 'zooming in' towards the posterior distributions sought. Both methods can potentially reduce the number of simulations required significantly, but they may be harder to parallelise than basic rejection-ABC. In addition, they may require more work to optimise: Defining *how* to move towards the best-fitting parameters can be difficult, and if done incorrectly, algorithms may "get stuck" in the wrong areas of the parameter space. However, SMC-ABC is less vulnerable to these problems, and may be worth trying with IBMs; Thiele et al. (2014) provide some introductory examples.

Improvements to the estimation of the posterior parameter distributions in Step 5 may also be possible. Known as "regression methods" (Beaumont et al., 2002; Blum and François, 2010), these techniques correct for the mismatch between the empirical data and the model outputs in the accepted runs. Inevitably, some accepted runs are going to be closer to the empirical data than others, but in basic rejection-ABC, all these runs contribute equally to the estimate of the posterior distributions. Regression methods attempt to correct for this anomaly by analysing the relationship between the parameter values and the summary statistics in the accepted runs, and then correcting parameter values accordingly. The *abc* package implements various ways of doing this correction (Csillery et al., 2012), but may produce unreliable results if some of the empirical data lies far outside the range of model outputs, as can happen with IBMs.

Our hope in providing this introduction to ABC is to persuade more ecological modellers to try it. Although the literature on ABC is large and growing, it is still mainly applied to population genetics problems. This means that it is still uncertain whether ABC's existing conventions and innovations are optimal for IBMs. For instance, choosing appropriate summary statistics is a field in its own right—if the available empirical data is summarised incorrectly, ABC's posteriors may be biased, or require many more simulations to get right (Blum et al., 2013); no general strategy can yet be advocated for IBMs. Other questions include whether ABC's typical distance measure (Eq. (4)) is the best choice for the time series data sometimes available in ecological applications, and how best to handle stochasticity. When we do simulation runs, we try each parameter combination once, but for some models, averaging over multiple runs with the same parameter values might be better. Finally, whether advanced techniques such as MCMC-ABC, SMC-ABC and the "regression correction" will prove workable with IBMs in practice is yet to be investigated. IBMs often have many more parameters than typical population genetics models, and different kinds of dependencies between them—only by trying things out, with lots of different IBMs, can general strategies be developed.

4. Conclusion

Science is a method of acquiring knowledge, and IBMs can be used to represent existing knowledge in ways that can be used to predict what will happen to individuals and populations in defined landscapes. Physiological ecology contributes the knowledge of how individuals acquire and expend energy, while behavioural ecology covers the factors that affect foraging, competition and social coexistence. Integrating these insights into IBMs allows us to link the levels from individuals to populations better than has been possible before. Even so, open questions remain. For example, at the individual level, there are still controversies about how animals distribute energy between physiological processes, and what they do when there is energy shortfall. At the population level, we are only just beginning to integrate social structures like dominance hierarchies into practical simulation models.

Thus, building realistic IBMs still requires expert judgement, and extensive testing against empirical data. Approximate Bayesian Computation, or ABC, is one possible approach to making this process more quantitative and transparent. Whereas the current state of the art, 'pattern-oriented modelling', or POM, is essentially a verbal protocol, ABC offers a statistically rigorous approach to model fitting and model comparison. However, ABC is fully compatible with the basic philosophy behind POM: That multiple empirical patterns, at multiple levels of organisation, should be used to build flexible, mechanistic models, that truly capture the fundamental aspects of the species and situations under consideration.

Although there are some challenges in implementing ABC for IBMs – most notably, the computing power required to evaluate models with long running times – the promise is considerable. ABC provides approximate posterior distributions of a model's parameters given data. As illustrated by our example, these posteriors can then be used as priors for further studies, and they can reveal which parameters are correlated or underconstrained. They can also be used to show the uncertainties that exist in a model's future predictions. Equally importantly, the power of ABC goes beyond parameter estimation–it can also be used to compare structurally different models, while automatically compensating for differences in model complexity.

Finally, we believe that perhaps one of the greatest advantages of ABC lies in its unifying language. Current efforts to parametrise IBMs, and to quantify their uncertainties, are often highly model dependent, with different types of results and plots provided in different studies. In contrast, ABC offers a set of conventional ways to report priors, posteriors, credible intervals and Bayes factors, and to do posterior checks and cross-validation and to calculate coverage. This should make model-fitting procedures more transparent. In addition, a basic understanding of ABC offers an entry point to the more sophisticated model-fitting alternatives that are available in the statistics literature. In sum, we believe ABC has much to offer when it comes to building, calibrating and evaluating realistic IBMs.

Acknowledgments

This work was supported by the Natural Environmental Resource Council [grant number NE/K006282/1, awarded to RM Sibly, M Beaumont, A Meade, PJ van Leeuwen and NK Nichols]. This work used the ARCHER UK National Supercomputing Service (http://www.archer.ac.uk). We are very grateful to Volker Grimm and two anonymous referees for constructive comments.

References

Beaumont, M.A., 2010. Approximate Bayesian computation in evolution and ecology. Annu. Rev. Ecol., Evol. Syst. 41, 379–406.

Beaumont, M.A., Zhang, W., Balding, D.J., 2002. Approximate Bayesian computation in population genetics. Genetics 162, 2025–2035.

Begon, M., Townsend, C.R., Harper, J.L., 2006. Ecology: From Individuals to Ecosystems, fourth ed. Blackwell Publishing, Malden, MA.

Bertorelle, G., Benazzo, A., Mona, S., 2010. ABC as a flexible framework to estimate demography over space and time: some cons, many pros. Mol. Ecol. 19, 2609–2625.

Blum, M.G.B., François, O., 2010. Non-linear regression models for approximate Bayesian computation. Stat. Comput. 20, 63–73.

Blum, M.G.B., Nunes, M.A., Prangle, D., Sisson, S.A., 2013. A comparative review of dimension reduction methods in approximate Bayesian computation. Stat. Sci. 28, 189–208.

Brown, J.H., Sibly, R.M., 2012. The metabolic theory of ecology and its central equation. In: Sibly, R.M., Brown, J.H., Kodric-Brown, A. (Eds.), Metabolic Ecology: A Scaling Approach. Wiley-Blackwell, Oxford, pp. 21–33.

Clauss, M., Schwarm, A., Ortmann, S., Streich, W.J., Hummel, J., 2007. A case of non-scaling in mammalian physiology? Body size, digestive capacity, food intake, and ingesta passage in mammalian herbivores. Comp. Biochem. Physiol., A: Mol. Integr. Physiol. 148, 249–265.

Csilléry, K., Blum, M.G.B., Gaggiotti, O.E., François, O., 2010. Approximate Bayesian computation (ABC) in practice. Trends Ecol. Evol. 25, 410–418.

Csillery, K., François, O., Blum, M.G.B., 2012. abc: An R package for approximate Bayesian computation (ABC). Methods Ecol. Evol. 3, 475–479.

Dalkvist, T., Sibly, R.M., Topping, C.J., 2011. How predation and landscape fragmentation affect vole population dynamics. PLoS ONE 6, e22834.

Davies, N.B., Krebs, J.R., West, S.A., 2012. An Introduction to Behavioural Ecology, fourth ed. Wiley-Blackwell, Oxford, UK.

DeAngelis, D.L., Mooij, W.M., 2005. Individual-based modeling of ecological and evolutionary processes. Annu. Rev. Ecol., Evol. Syst. 36, 147–168.

Evans, M.R., Bithell, M., Cornell, S.J., Dall, S.R.X., Diaz, S., Emmott, S., Ernande, B., Grimm, V., Hodgson, D.J., Lewis, S.L., Mace, G.M., Morecroft, M., Moustakas, A., Murphy, E., Newbold, T., Norris, K.J., Petchey, O., Smith, M., Travis, J.M.J., Benton, T.G., 2013. Predictive systems ecology. Proc. R. Soc. B: Biol. Sci. 280, 20131452.

Galic, N., Forbes, V., 2014. Ecological models in ecotoxicology and ecological risk assessment: an introduction to the special section. Environ. Toxicol. Chem. 33, 1446–1448.

Glazier, D.S., 2005. Beyond the '3/4-power law': variation in the intra- and interspecific scaling of metabolic rate in animals. Biol. Rev. 80, 611–662.

Glazier, D.S., 2008. Resource allocation patterns. In: Rauw, W.M. (Ed.), Resource Allocation Theory Applied to Farm Animal Production. CABI Publishing, Wallingford, UK, pp. 22–43.

Gosler, A.G., Greenwood, J.J.D., Perrins, C., 1995. Predation risk and the cost of being fat. Nature 377, 621–623.

Grimm, V., Railsback, S.F., 2005. Individual-based Modeling and Ecology. Princeton University Press, Princeton, NJ.

Grimm, V., Railsback, S.F., 2012. Pattern-oriented modelling: a 'multi-scope' for predictive systems ecology. Philos. Trans. R. Soc. B: Biol. Sci. 367, 298–310.

Gunadi, B., Blount, C., Edwards, C.A., 2002. The growth and fecundity of *Eisenia fetida* (Savigny) in cattle solids pre-composted for different periods. Pedobiologia 46, 15–23.

Gunadi, B., Edwards, C.A., 2003. The effects of multiple applications of different organic wastes on the growth, fecundity and survival of *Eisenia fetida* (Savigny) (Lumbricidae). Pedobiologia 47, 321–329.

Hartig, F., Calabrese, J.M., Reineking, B., Wiegand, T., Huth, A., 2011. Statistical inference for stochastic simulation models—theory and application. Ecol. Lett. 14, 816–827.

Hartig, F., Dislich, C., Wiegand, T., Huth, A., 2014. Technical note: approximate Bayesian parameterization of a process-based tropical forest model. Biogeosciences 11, 1261–1272.

Hartman, K.J., Kitchell, J.F., 2008. Bioenergetics modeling: progress since the 1992 symposium. Trans. Am. Fish. Soc. 137, 216–223.

Hendriks, A.J., 1999. Allometric scaling of rate, age and density parameters in ecological models. Oikos 86, 293–310.

Holling, C.S., 1959. The components of predation as revealed by a study of small-mammal predation of the European pine sawfly. Can. Entomol. 91, 293–320.

Hou, C., 2014. Increasing energetic cost of biosynthesis during growth makes refeeding deleterious. Am. Natur. 184, 233–247.

Johnston, A.S.A., Hodson, M.E., Thorbek, P., Alvarez, T., Sibly, R.M., 2014. An energy budget agent-based model of earthworm populations and its application to study the effects of pesticides. Ecol. Modell. 280, 5–17.

Karasov, W.H., Martinez del Rio, C., 2007. Physiological Ecology. Princeton University Press, Princeton, NJ and Oxford, UK.

Kaspari, M., 2012. Stoiciometry. In: Sibly, R.M., Brown, J.H., Kodric-Brown, A. (Eds.), Metabolic Ecology: A Scaling Approach. Wiley-Blackwell, Oxford, UK, pp. 14–47.

Kass, R.E., Raftery, A.E., 1995. Bayes factors. J. Am. Stat. Assoc. 90, 773–795.

Kerkhoff, A.J., 2012. Modeling metazoan growth and ontogeny. In: Sibly, R.M., Brown, J.H., Kodric-Brown, A. (Eds.), Metabolic Ecology: A Scaling Approach. Wiley-Blackwell, Oxford, UK.

Kooijman, S.A.L.M., 2010. Dynamic Energy Budget Theory, third ed. Cambridge University Press, Cambridge, UK.

Krebs, C.J., 2009. Ecology, sixth ed. Benjamin Cummings, San Francisco, CA.

Kulakowska, K.A., Kulakowski, T.M., Inglis, I.R., Smith, G.C., Haynes, P.J., Prosser, P., Thorbek, P., Sibly, R.M., 2014. Using an individual-based model to select among alternative foraging strategies of woodpigeons: Data support a memory-based model with a flocking mechanism. Ecol. Modell. 280, 89–101.

Lind, J., Jakobsson, S., Kullberg, C., 2010. Impaired predator evasion in the life history of birds: behavioral and physiological adaptations to reduced flight ability. In: Thompson, C.F. (Ed.), Current Ornithology, vol. 17. Springer, New York, pp. 1–30.

Lindstedt, S.L., Schaeffer, P.J., 2002. Use of allometry in predicting anatomical and physiological parameters of mammals. Lab. Anim. 36, 1–19.

Liu, C., Sibly, R.M., Grimm, V., Thorbek, P., 2013. Linking pesticide exposure and spatial dynamics: an individual-based model of wood mouse (*Apodemus sylvaticus*) populations in agricultural landscapes. Ecol. Modell. 248, 92–102.

Marjoram, P., Molitor, J., Plagnol, V., Tavaré, S., 2003. Markov chain Monte Carlo without likelihoods. Proc. Natl. Acad. Sci. U.S.A. 100, 15238–15324.

Martin, B.T., Jager, T., Nisbet, R.M., Preuss, T.G., Grimm, V., 2013. Predicting population dynamics from the properties of individuals: a cross-level test of dynamic energy budget theory. Am. Natur. 181, 506–519.

Martin, B.T., Zimmer, E.I., Grimm, V., Jager, T., 2012. Dynamic energy budget theory meets individual-based modelling: a generic and accessible implementation. Methods Ecol. Evol. 3, 445–449.

Martínez, I., Wiegand, T., Camarero, J.J., Batllori, E., Gutiérrez, E., 2011. Disentangling the formation of contrasting tree-line physiognomies combining model selection and Bayesian parameterization for simulation models. Am. Natur. 177, E136–E152.

Monroy, F., Aira, M., Domínguez, J., Velando, A., 2006. Seasonal population dynamics of *Eisenia fetida* (Savigny, 1826) (Oligochaeta, Lumbricidae) in the field. Comptes Rendus Biol. 329, 912–915.

Moses, M.E., Hou, C., Woodruff, W.H., West, G.B., Nekola, J.C., Zuo, W., Brown, J.H., 2008. Revisiting a model of ontogenetic growth: estimating model parameters from theory and data. Am. Natur. 171, 632–645.

Nabe-Nielsen, J., Sibly, R.M., Forchhammer, M.C., Forbes, V.E., Topping, C.J., 2010. The efects of landscape modifications on the long-term persistence of animal populations. PLoS ONE 5, e8932.

Nabe-Nielsen, J., Sibly, R.M., Tougaard, J., Teilmann, J., Sveegaard, S., 2014. Effects of noise and by-catch on a Danish harbour porpoise population. Ecol. Modell. 272, 242–251.

Peters, R.H., 1983. The Ecological Implications of Body Size. Cambridge University Press, Cambridge, UK.

Petit, O., Gautrais, J., Leca, J.B., Theraulaz, G., Deneubourg, J.-L., 2009. Collective decision making in white-faced capuchin monkeys. Proc. R. Soc. B: Biol. Sci. 1672, 3495–3503.

Pond, C.M., 1978. Morphological aspects and ecological and mechanical consequences of fat deposition in wild vertebrates. Ann. Rev. Ecol. Syst. 9, 519–570.

Prangle, D., Blum, M.G.B., Popovic, G., Sisson, S.A., 2013. Diagnostic tools for approximate Bayesian computation using the coverage property. Aust. N. Z. J. Stat. 56, 309–329.

Price, C.A., Weitz, J.S., Savage, V.M., Stegen, J., Clarke, A., Coomes, D.A., Dodds, P.S., Etienne, R.S., Kerkhoff, A.J., McCulloh, K., Niklas, K.J., Olff, H., Swenson, N.G., 2012. Testing the metabolic theory of ecology. Ecol. Lett. 15, 1465–1474.

Pritchard, J.K., Seielstad, M.T., Perez-Lezaun, A., Feldman, M.W., 1999. Population growth of human Y chromosomes: a study of Y chromosome microsatellites. Mol. Biol. Evol. 16, 1791–1798.

Reinecke, A.J., Viljoen, S.A., 1990. The influence of feeding patterns on growth and reproduction of the vermicomposting earthworm *Eisenia fetida* (Oligochaeta). Biol. Fertil. Soils 10, 184–187.

Ricklefs, R.E., Miller, G.L., 2000. Ecology, fourth ed. W.H. Freeman and Co., New York, NY.

Sibly, R.M., 2002. Life history theory. In: Pagel, M. (Ed.), Encyclopedia of Evolution. Oxford University Press, Oxford, pp. 623–627.

Sibly, R.M., Calow, P., 1986. Physiological Ecology of Animals. Blackwell Scientific Publications, Oxford, UK.

Sibly, R.M., Grimm, V., Martin, B.T., Johnston, A.S.A., Kulakowska, K., Topping, C.J., Calow, P., Nabe-Nielsen, J., Thorbek, P., DeAngelis, D.L., 2013. Representing the acquisition and use of energy by individuals in agent-based models of animal populations. Methods Ecol. Evol. 4, 151–161.

Sih, A., Cote, J., Evans, M., Fogarty, S., Pruitt, J., 2012. Ecological implications of behavioural syndromes. Ecol. Lett. 15, 278–289.

Simpson, S.J., Sibly, R.M., Lee, K.P., Behmer, S.T., Raubenheimer, D., 2004. Optimal foraging when regulating intake of multiple nutrients. Anim. Behav. 68, 1299–1311.

Sinclair, A.R.E., 1989. Population regulation in animals. In: Cherrett, J.M. (Ed.), Ecological Concepts. Blackwell Scientific, Oxford, pp. 197–241.

Sisson, S.A., Fan, Y., Tanaka, M.A., 2007. Sequential Monte Carlo without likelihoods. Proc. Natl. Acad. Sci. U.S.A. 104, 1760–1765.

Stearns, S.C., 1992. The Evolution of Life Histories. Oxford University Press, Oxford, UK.

Stillman, R.A., Goss-Custard, J.D., 2010. Individual-based ecology of coastal birds. Biol. Rev. 85, 413–434.

Stillman, R.A., Railsback, S.F., Giske, J., Berger, U., Grimm, V., 2015. Making predictions in a changing world: the benefits of individual-based ecology. BioScience 65, 140–150.

Tavaré, S., Balding, D.J., Griffiths, R.C., Donnelly, P., 1997. Inferring coalescence times from DNA sequence data. Genetics 145, 505–518.

Thiele, J.C., Kurth, W., Grimm, V., 2012. RNetLogo: an R package for running and exploring individual-based models implemented in NetLogo. Methods Ecol. Evol. 3, 480–483.

Thiele, J.C., Kurth, W., Grimm, V., 2014. Facilitating parameter estimation and sensitivity analysis of agent-based models: a cookbook using NetLogo and R. J. Artif. Soc. Soc. Simul. 17, 11.

Toni, T., Welch, D., Strelkowa, N., Ipsen, A., Stumpf, M.P.H., 2009. Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. J. R. Soc. Interf. 6, 187–202.

Topping, C.J., Dalkvist, T., Grimm, V., 2012. Post-hoc pattern-oriented testing and tuning of an existing large model: Lessons from the field vole. PLoS ONE 7, e45872.

van der Vaart, E., Beaumont, M.A., Johnston, A., Sibly, R.M., 2015a. Calibration and evaluation of individual-based models using approximate Bayesian computation. Ecol. Modell. 312, 182–190.

van der Vaart, E., Johnston, A., Sibly, R.M., 2015b. Linking Levels—Runs. Figshare, 1494757, doi: http://dx.doi.org/10.6084/m9.figshare.

van der Vaart, E., Johnston, A.S.A., Sibly, R.M., 2015c. Linking Levels—Code. Figshare, 1494754, doi: http://dx.doi.org/10.6084/m9.figshare.

Wilensky, U., 1999. NetLogo. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL, ⟨http://ccl.northwestern.edu/netlogo/⟩.

Witter, M.S., Cuthill, I.C., 1993. The ecological costs of avian fat storage. Philos. Trans. R. Soc. London Ser. B: Biol. Sci. 340, 73–92.